(2)

SECURITY CLASSIFICATION OF THIS PAGE

AD-A215 241

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1b. RESTRICTIVE MARKINGS

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 89-1454

6a. NAME OF PERFORMING ORGANIZATION
Marquette University
Department of Electrical

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
AFOSR

6c. ADDRESS *(City, State, and ZIP Code)*
Engineering
Milwaukee, Wisconsin 53233

7b. ADDRESS *(City, State, and ZIP Code)*
BLDG 410
BAFB DC 20332-6448

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR 78-3600

8c. ADDRESS *(City, State, and ZIP Code)*
BLDG 410
BAFB DC 20332-6448

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
61102F	2304	A6	

11. TITLE *(Include Security Classification)*
THE ENHANCEMENT OF SPEECH INTELLIGIBILITY IN WIDEBAND NOISE.

12. PERSONAL AUTHOR(S)
Russell J. Niederjohn

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM ______ TO ______	January 23, 1980	18

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

DTIC
ELECTE
NOV 30 1989
S D
D

89 11 .. 073

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL

22b. TELEPHONE *(Include Area Code)*
767-5025

22c. OFFICE SYMBOL
NM

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

AFOSR-TR- 89-1454

FINAL REPORT

AIR FORCE
OFFICE OF SCIENTIFIC RESEARCH

(June 1, 1978 - November 31, 1979)

Grant Title: The Enhancement of Speech Intelligibility in Wideband Noise

Grant Number: AFOSR-78-3600

Awarded To: Marquette University

Amount of Grant: $9,999

Principal Investigator: Dr. Russell J. Niederjohn

DATE: January 23, 1980

I. INTRODUCTION

The understanding of speech contaminated by the presence of noise is an important consideration in many practical communication situations. The research described in this report details the results of two related studies involving speech mixed with additive white noise. The first study concerns the effects of high-pass and low-pass filtering upon the intelligibility of speech in noise. This study has provided basic perceptual data concerning the speech in noise situation. The results of this study are described in detail in a report (1) (abstract attached) and in a manuscript recently submitted for publication in the IEEE Transactions on Acoustics, Speech and Signal Processing (2) (abstract also attached). An overview of the significant results of this study are contained in Section II of this report.

The second study concerns the development of a method for processing speech mixed with noise to result in an enhancement of its intelligibility. While several processing methods were implemented, the method explored in greatest detail is based upon a combination of a spectral subtraction method and a pitch tracking method. The final intelligibility results obtained using the method did not show that it resulted in an intelligibility enhancement over unprocessed speech in noise. However, the results are very enlightening regarding the speech in noise problem. In addition, it is this author's belief that the basic processing method has significant potential for enhancing the intelligibility of speech in noise. Further research will need to be performed to improve upon the intelligibility results using this method. An overview of the significant results of this speech processing work are contained in Section III of this report. Greater detail can be found in a recently published report (3) (abstract attached).

The research performed under this grant was intended to consider only the case where the noise is at the speaker. The case of noise at the listener has been considered in previous work by the author and others (4-9). In addition, the work performed under this grant was restricted to the case of additive white noise. However, the processing method considered (Section III) was purposely designed so that it could be extended to the case of non-white noise in a fairly straightforward manner.

II. THE EFFECT OF LOW-PASS AND HIGH-PASS FILTERING UPON THE INTELLIGIBILITY OF SPEECH IN NOISE

The research performed in this part of the study is intended to provide a fundamental background in the speech in noise situation for the case of noise at the speaker. In particular, this part of the study is concerned with the effects of high-pass and low-pass filtering upon the intelligibility of speech in white noise. The goal is to provide basic perceptual information about the intelligibility degradation of speech mixed with additive white noise and to provide a quantitative measure of the contribution to intelligibility of various frequency ranges of the speech spectrum as a function of speech-to-noise ratio. The results of this study will be of a significant interest to researchers pursuing methods for the intelligibility enhancement of speech in noise.

The experimental procedures, speech materials, and speech and noise measurement methods used are described in detail in a recent report (1). The main intelligibility results are shown in Figures 1 and 2 (reproduced from (2)) for low-pass and for high-pass filtered speech respectively. Several conclusions concerning these results and their implication for the speech in noise situation are presented in a recent paper (2). For the case of white noise under the testing conditions and over the speech and noise levels tested in this work, the main conclusions of this study are summarized as follows: (1) No low-pass or high-pass filter enhances the intelligibility of speech in noise relative to unfiltered speech in noise; (2) The intelligibility of high-pass filtered speech decreases more rapidly with increasing cutoff frequency as the signal-to-noise ratio is lowered; (3) There is relatively little increase in speech intelligibility using low-pass filters having a cutoff frequency greater than about 1.5 kHz; (4) The listening level preferred by subjects attempting to maximize their perceived intelligibility of high-pass filtered speech in noise is determined almost solely from the speech level, indepedent of speech-to-noise ratio. A somewhat similar, although less consistent, effect is observed for the preferred listening level for speech processed by low-pass filters; (5) It appears, from the data collected in this study in comparison to the data collected in an earlier study (4), that listeners are able to optimize their actual speech intelligibility by optimizing their perceived speech intelligibility.

An examination of the articulation index (AI) and its relationship with the experimental data collected in this study has provided a verification of the data and has provided an additional validation of the articulation index method. The agreement of the curve of speech intelligibility versus AI (from Figure 4 in (2)) with the expected curve (from (15)) indicates not only the reliability of the data in Figures 1 and 2 but also it indicates that the AI method is applicable to systems with variable gain. The authors have not found any previous reference to the use of the AI with variable gain systems.

From the results of this study, it is possible to predict, using the AI method, the expected intelligibility for any signal-to-noise ratio (over the range of signal-to-noise ratios tested) and any high-pass or low-pass filter (over the range of filters tested). It is reasonable to expect that the results of this study are additionally applicable over a wide range signal-to-noise ratios and over a wide range of filter configurations.

III. PROCESSING SPEECH TO ENHANCE ITS INTELLIGIBILITY IN NOISE

The objective of the second part of this research effort was to evaluate several frequency domain processing techniques with the goal of determining a method to enhance the intelligibility of speech in noise. Through this research effort several frequency domain techniques were evaluated. The techniques evaluated included several variations of methods based upon (a) spectral subtraction, (b) pitch tracking, (c) minimum mean-square-error filtering, and (d) INTEL. This later technique was developed by Weiss, et. al. under Air Force support and is documented in two reports (10-11). A report by the author provides some additional insite into and documentation of the INTEL algorithm (12).

The method explored in greatest detail is based upon a combination of a spectral subtraction technique and a pitch tracking based technique. In Figure 3 (from (3)) is shown a flow diagram for the most current implementation of this processing technique. The algorithm involves one forward FFT and one inverse FFT. All processing is done on the frequency domain signal as shown in Figure 3.

Spectral subtraction is used in the method to remove a significant amount of base-line noise while pitch processing is used to remove noise between pitch harmonics. The method incorporates a pitch tracking method based upon a variation in the harmonic sum spectrum proposed by Noll (13). In addition, it incorporates a voiced/unvoiced detector (for noisy speech) and different processing for voiced and for unvoiced speech. The entire algorithm has been implemented on a PDP-11/34 computer system using principally FORTRAN program coding.

An evaluation of the processing method was carried out using fifteen University students. Tests were conducted at five signal-to-noise ratios using a testing procedure and test material from Nye and Gaitenby (14). Overall the test results showed that the processing method did not result in an intelligibility enhancement at any signal-to-noise ratio. This discouraging result is not unlike the results other researchers have obtained with other speech-in-noise processing methods. A speech processing method which enhances the intelligibility of speech in noise has yet to be demonstrated. Also, like other speech-in-noise processing methods, this method does result in an enhancement of the so called "listenability." In evaluating the results from speech processing systems it is often difficult to subjectively separate listenability from intelligibility.

IV. CONCLUSIONS

Through this study, the author of this report and the graduate students who worked with him accumulated a significantly better understanding of the speech in noise problem. Some of this understanding has been documented in a paper recently submitted for publication (2). The other research performed under this grant has been documented in two recent reports (1,3).

Through the evaluation of the intelligibility of low-pass and high-pass filtered speech in noise (first part of this study) some fundamental quantitative data have been obtained for the speech in noise problem. Even though the goal of determining a speech processing method to result in an enhancement of the intelligibility of speech in noise was not met (second part of this study), the results obtained have provided considerable insite into speech in noise processing methods. Further, it is this author's belief that additional work with this algorithm has a significant potential for resulting in a method for enhancing the intelligibility of speech in noise.

The significance of several speech processing problems treated in this work should not be overlooked. In particular, the pitch tracking method developed was qualitatively evaluated and found to be very accurate in spite of the additive noise. In addition, the voiced/unvoiced detector developed for use with speech in noise was also qualitatively determined to be quite accurate. Finally, the adaptive technique used to automatically adjust the process as a function of signal-to-noise ratio, and the use of separate processing for voiced and for unvoiced speech are significant contributions of this research.

REFERENCES

(1) David G. Mliner, "The measurement and prediction of the effect of low-pass and high-pass filtering upon the intelligibility of speech in white noise," Department of Electrical Engineering, Marquette University, Technical Report No. EE-79-2, 1979 (M.S. Thesis).

(2) Russell J. Niederjohn and David G. Mliner, "The effects of high-pass and low-pass filtering upon the intelligibility of speech in white noise," IEEE Transactions on Acoustics, Speech, and Signal Processing, submitted for publication.

(3) Barry J. Sullivan, "Intelligibility enhancement of speech in wide-band random noise by pitch processing and spectral subtraction," Department of Electrical Engineering, Marquette University, Technique Report No. EE-79-1, 1979 (M.S. Thesis).

(4) R. J. Niederjohn and J. H. Grotelueschen, "The enhancement of speech intelligibility in high noise levels by high-pass filtering followed by rapid amplitude compression," IEEE Trans. Acoust., Speech, and Signal Processing, ASSP-24, 277-282, 1976.

(5) R. J. Niederjohn and J. H. Grotelueschen, "Speech intelligibility enhancement in a power generating noise environment," IEEE Trans. Acoust. Speech, and Signal Processing, ASSP-26, 378-380, 1978.

(6) I. B. Thomas and R. J. Niederjohn, "The intelligibility of filtered clipped speech in noise," J. Aud. Eng. Soc., 18, 299-303, 1970.

(7) I. B. Thomas and W. J. Ohley, "Intelligibility enhancement through spectral weighting," 1972 Conf. on Speech Communication and Processing, 360-363, 1972.

(8) I. Pollack and J. M. Pickett, "Intelligibility of peak-clipped speech at high noise levels," J. Acoust. Soc. Am., 31, 14-16, 1959.

(9) R. J. Niederjohn, "A comparison of three recently reported methods for processing speech for the enhancement of speech intelligibility in high noise levels," Proc. 1976 Midwest Symp. on Circuits and Systems, 325-331, 1976.

(10) M. R. Weiss, E. Aschkenasy, and T. W. Parsons, "Study and development of the INTEL technique for improving speech intelligibility," Technical Report No. RADC-TR-75-108, Rome Air Development Center, Griffiss Air Force Base, New York, April, 1975.

(11) M. R. Weiss and E. Aschkenasy, "Automatic detection and enhancement of speech signals," Technical Report No. RADC-TR-75-77, Rome Air Development Center, Griffiss Air Force Base, New York, March, 1975.

(12) R. J Niederjohn and R. A. Curtis, "A system for computer speech processing and its application for the enhancement of speech intelligibility in noise," Technical Report No. RADC-TR-77-310, Rome Air Development Center, Griffiss Air Force Base, New York, 1977.

(13) A. M. Noll, "Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum, and a maximum liklihood estimate," in MRI Symp. Proc. Polytechnic Inst. Brooklyn, 1970, Brooklyn, N.Y.: Polytechnic Press, Vol. XIX, pp. 779-797.

(14) P. W. Nye and J. H. Gaitenby, "The intelligibility of monosyllabic words in short, syntactically normal sentences," Haskins Lab., New Haven, CT, Status Rept. on Speech Res., SR-37/38, pp. 169-190, 1974.

(15) American National Standard Method for the Calculation of the Articulation Index, ANSI 53.5 - 1969.

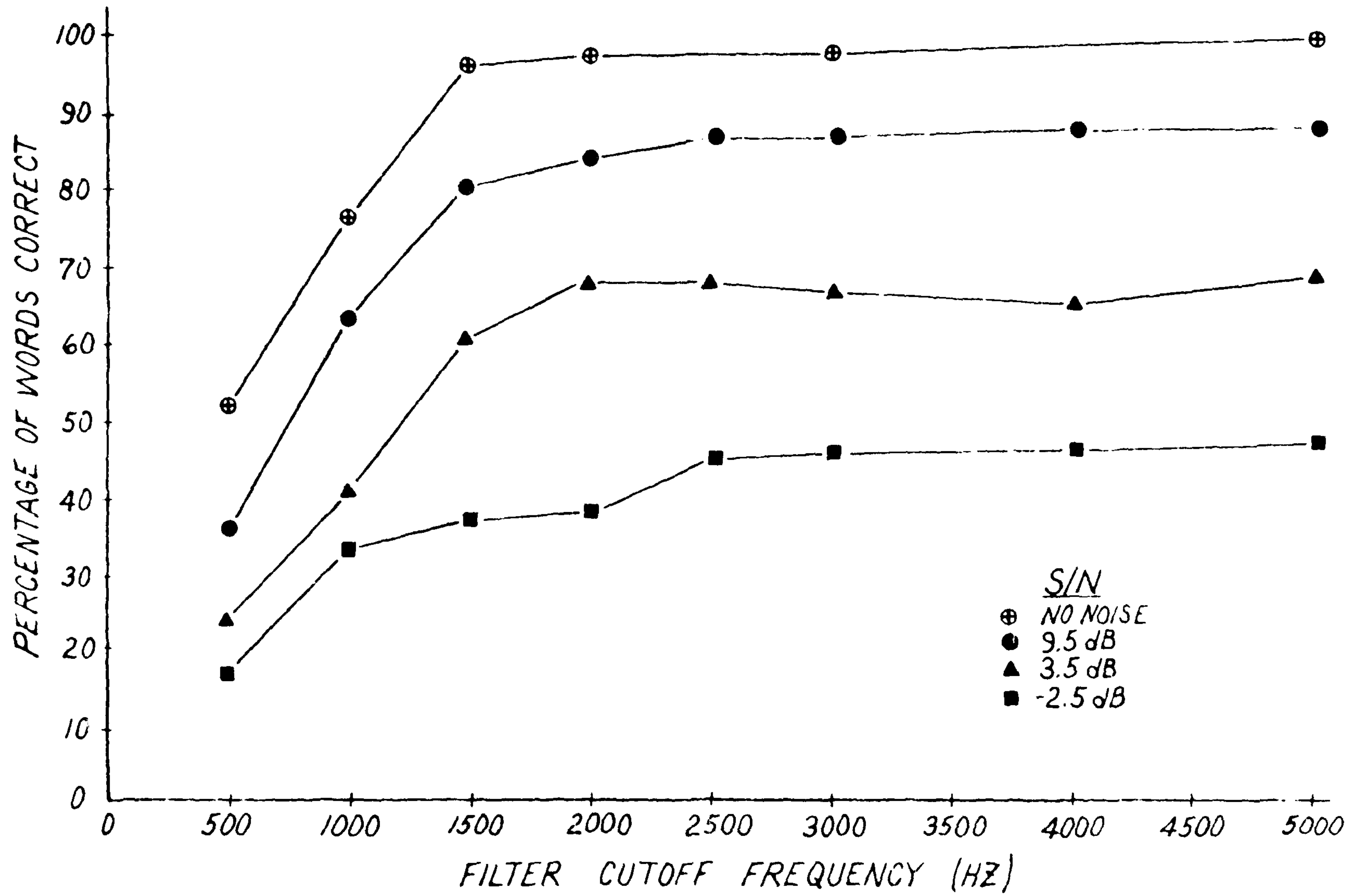

Figure 1 - Intelligibility results for low-pass filtered speech at four signal-to-noise ratios.

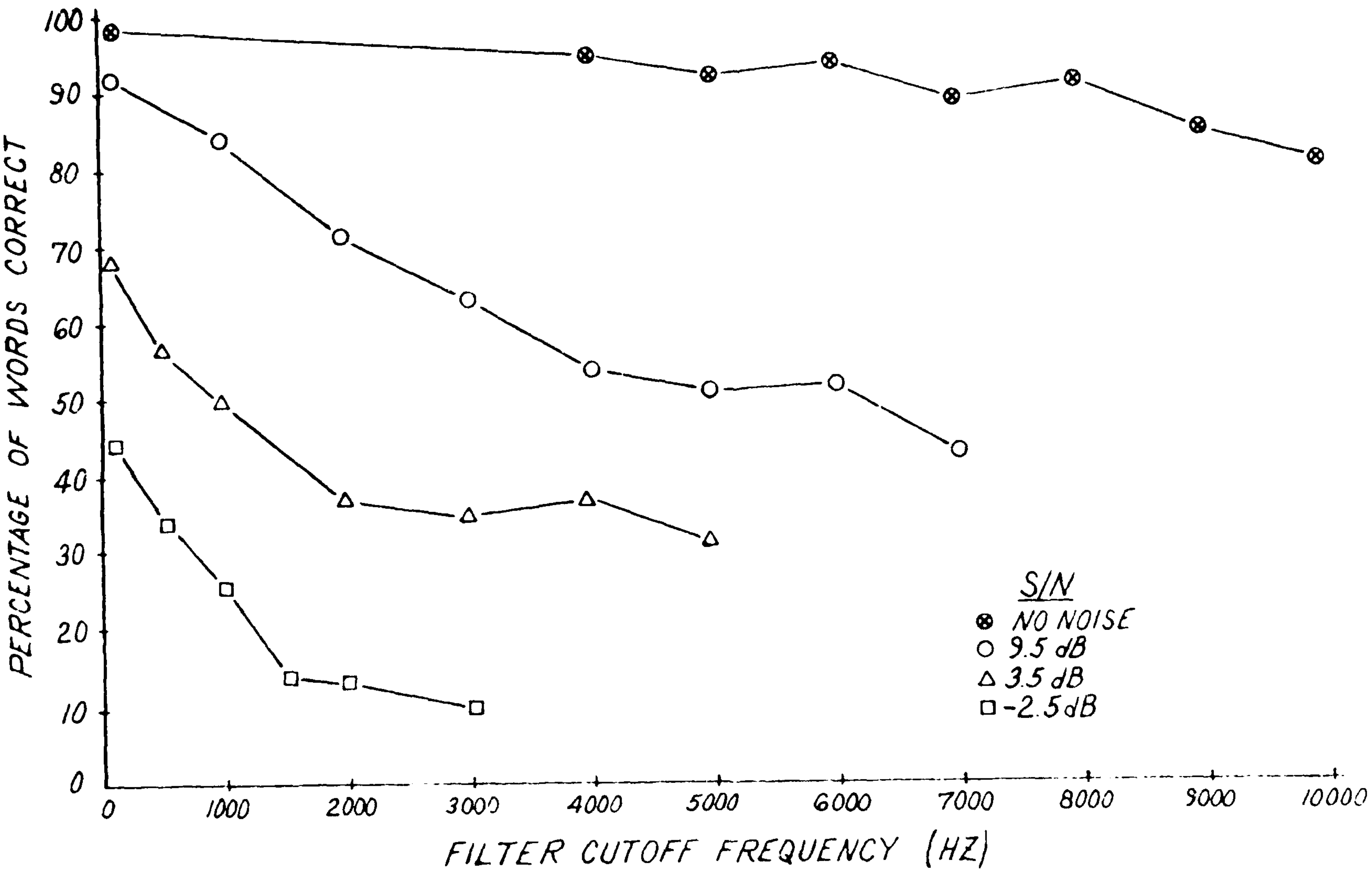

Figure 2 - Intelligibility results for high-pass filtered speech at four signal-to-noise ratios.

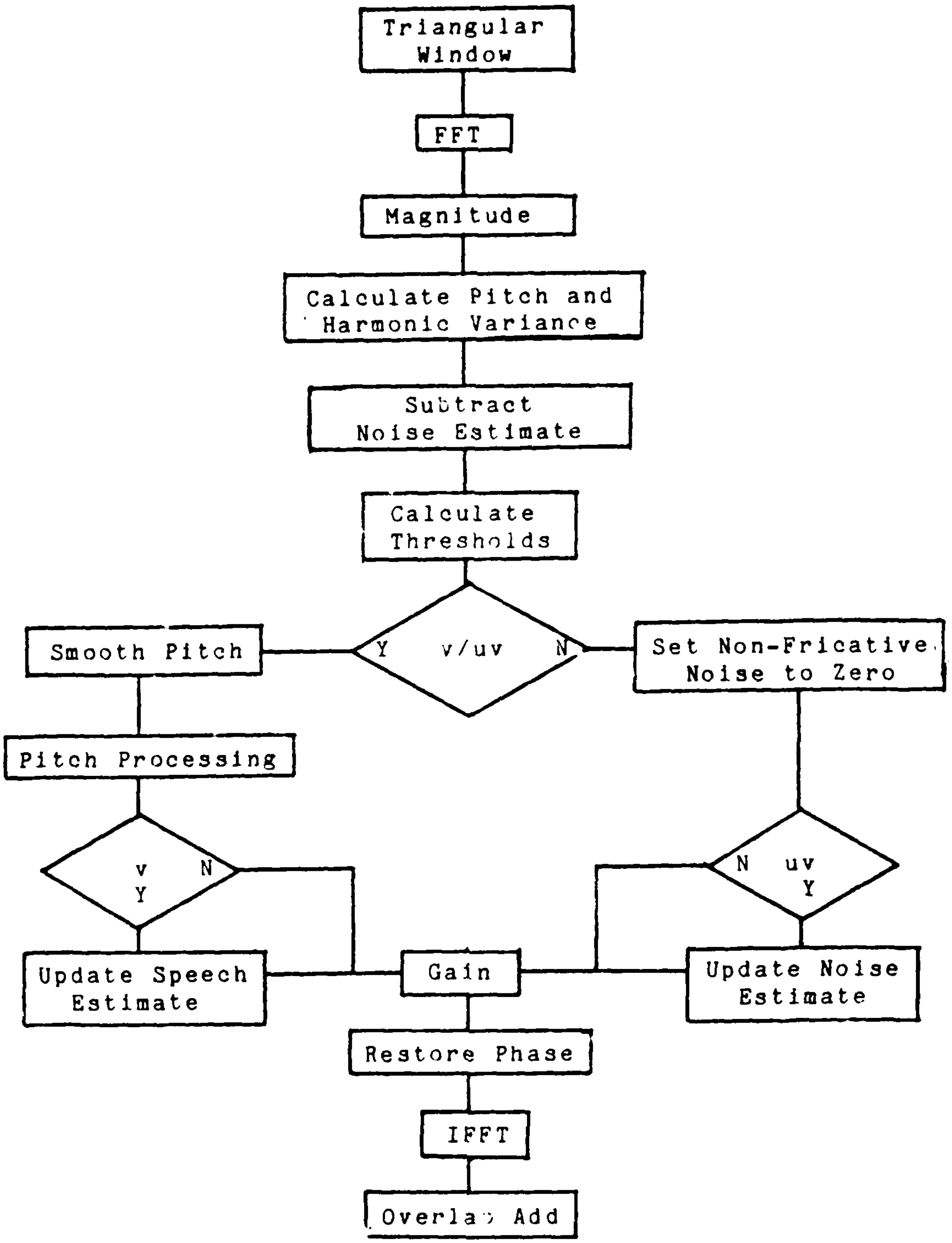

Figure 3 - Flow diagram of the algorithm used for processing speech to result in an enhancement of its intelligibility in noise.

APPENDIX

Bibliography of papers and reports describing the research performed under this grant. An abstract of each is attached as follows:

Appendix A: R. J. Niederjohn and D. G. Mliner, "The effects of high-pass and low-pass filtering upon the intelligibility of speech in white noise," IEEE Transactions on Acoustics, Speech, and Signal Processing, submitted for publication.

Appendix B: D. G. Mliner, "The measurement and prediction of the effect of low-pass and high-pass filtering upon the intelligibility of speech in white noise," Department of Electrical Engineering, Marquette University, Technical Report No. EE-79-2, 1979, (M.S. Thesis).

Appendix C: B. J. Sullivan, "Intelligibility enhancement of speech in wide-band random noise by pitch processing and spectral subtraction," Department of Electrical Engineering, Marquette University, Technical Report No. EE-79-1, 1979 (M.S. Thesis).

APPENDIX A

Abstract of a manusctipt recently submitted to the IEEE Transactions on Acoustics, Speech, and Signal Processing.

THE EFFECTS OF HIGH-PASS AND LOW-PASS FILTERING UPON THE INTELLIGIBILITY OF SPEECH IN WHITE NOISE

BY

Russell J. Niederjohn and David G. Mliner*

Electrical Engineering Department
Marquette University
Milwaukee, Wisconsin 53233

ABSTRACT

This paper presents the results of a study of the effects of high-pass and low-pass filtering upon the intelligibility of speech mixed with additive white noise. Experiments are described wherein the intelligibility of high-pass and low-pass filtered speech, at four signal-to-noise ratios is measured with trained listeners. These perceptual intelligibility results are then interpreted to provide information concerning the contribution to intelligibility of various frequency ranges of the speech spectrum. Finally, the results are examined in realtion to the articulation index. The perceptual results provide good agreement with the articulation index method and provide an improved understanding of the speech in noise situation.

---

* Presently with Rockwell International, Anaheim, California 92803

APPENDIX B

REPORT* ABSTRACT

Enhancement of the intelligibility of speech in noise is a complex and unsolved problem. In order to gain some understanding of the speech in noise situation, the relative importance of various frequency ranges to the intelligibility of speech in noise is investigated in this thesis. Some conclusions which are drawn for the frequency range and S/N ratios used include:

(1) No high-pass or low-pass filter increases the intelligibility of speech in noise.

(2) Speech can be low-pass or high-pass filtered at 2500 Hz without significant loss of intelligibility.

(3) The intelligibility of speech decreases rapidly with increasing high-pass filter cutoff frequency.

(4) With speech in noise, listeners tend to set the level of the composite signal based upon the level of the speech, regardless of the S/N ratio.

The Articulation Index is calculated and validated for the experimental data gathered in this work.

* D. G. Mliner, "The measurement and prediction of the effect of low-pass and high-pass filtering upon the intelligibility of speech in white noise," Department of Electrical Engineering, Marquette University, Technical Report No. EE-79-2, 1979 (M.S. Thesis).

APPENDIX C

REPORT* ABSTRACT

The problem of speech rendered unintelligible by the addition of white noise is considered. A process with the goal of enhancing the intelligibility of the corrupted speech is developed and implemented on a digital computer. This implementation is evaluated by comparing the intelligibility of unprocessed speech with the intelligibility of processed speech at various noise levels. The results of the evaluation are presented and discussed.

* B. J. Sullivan, "Intelligibility enhancement of speech in wide-band random noise by pitch processing and spectral subtraction," Department of Electrical Engineering, Marquette University, Technical Report No. EE-79-1, 1979 (M.S. Thesis).